Sports Info Solutions Analytics Challenge

Stanford Sports Analytics Club

Matt Colón, Silas Morsink, Robbie Thompson, and Peter Gofen

Which Defensive Line Position is Most Important?

Our Approach:

Positions where variation in talent leads to larger variation in outcomes are more important

We investigate the relationship between talent and outcomes at two positional groups: interior and exterior defensive linemen. Using the principle above, we measure the relative importance of both groups, both in general and in specific contexts.

Under the data and time constraints, we opted for positional groupings as opposed to player-specific positions, and we used Madden ratings as a proxy for player talent. However, our methodology is flexible and can be executed with any defined positions or form of positional talent evaluation.

Top-Line Results

1. **Exterior defensive linemen** are the most valuable defensive line position group.
2. Regarding the distribution of talent between the two positions, exterior defensive linemen are slightly more valuable than interior defensive linemen.
3. The following two scenarios change our answer to Question 1:
    - As the probability of a run on a play increases:
        - (1) Exterior Defensive Linemen are still more important, but
        - (2) The difference in importance between Exterior and Interior Defensive Linemen becomes smaller
    - In a roster construction exclusively utilizing two Interior and two Exterior Defensive Linemen:
        - (1) The Interior Defensive Linemen position becomes most important
        - (2) Interior Defensive Linemen position is slightly more important than Exterior Defensive Linemen position

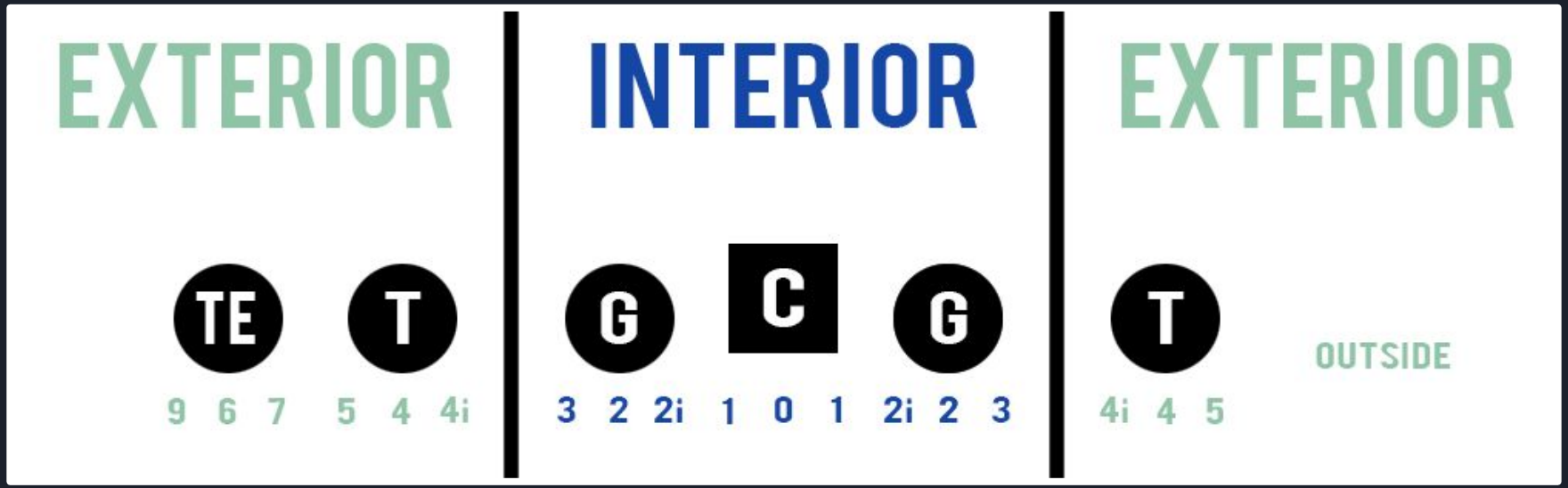
Methodology and Assumptions

Relating Variation Within Positions to Outcomes

Relating Variation Within **Positions** to Outcomes

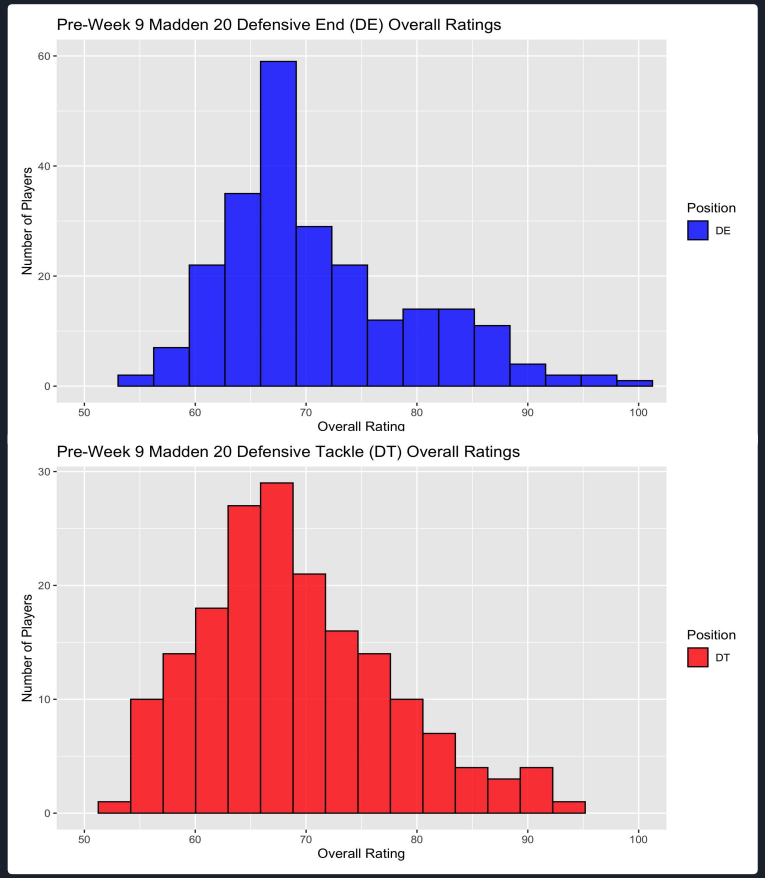
- To achieve stable results with our limited dataset, we break defensive line positioning into two categories: interior and exterior
- **Interior defensive linemen (IDL)** are defined as any on-ball defender lined up in a 3-technique or further inside
- **Exterior defensive linemen (EDL)** are all on-ball players outside of a 3-technique as well as off-ball defenders designated as being on the outside edge
    - This definition includes defensive ends who walk up in a 2-point stance as well as outside linebackers, who often function as a force player in run fits and rush the passer on the majority of dropbacks

Relating Variation Within **Positions** to Outcomes



Relating **Variation** Within Positions to Outcomes

- To measure positional talent, we use EA Sports' Madden NFL 20 Pre-Week 9 ratings
- We normalize player ratings to define comparable increases in talent across positions
  - i.e. Equating the difference between IDLs Fletcher Cox and DeForest Buckner to the difference between EDLs J.J. Watt and Joey Bosa
- While we use Madden ratings for our analysis, variation in any position-specific ratings metric could serve the same purpose



Relating Variation Within Positions to **Outcomes**

Outcome of interest: **Defensive-line adjusted, context-neutral EPA**

Defensive Line Adjustment

- We only want to reward/punish defensive linemen for plays that they affect
- We assume that defensive linemen influence which bucket a play falls into but have a negligible impact on variation in play outcome within a bucket

Context Neutrality

- We extrapolate away from the variation in EPA caused by game context (e.g. field position, down, and distance) by treating all plays within the same bucket the same

Relating Variation Within Positions to **Outcomes**

Through the defensive line and context neutrality adjustments, we reduce the noise in outcomes by extrapolating away from variation not related to defensive line actions

Right: Buckets and the Adjusted EPA for each bucket. This Adjusted EPA is our response variable—how we describe the outcome of each play

Bucket	Rush big loss	Rush small loss	Rush no gain	Rush small gain	Rush decent gain	Rush big gain	Screen	Short pass	Medium pass	Deep pass
Notes	-3 or fewer yards	-2 or -1 yards	0 or 1 yards	2 to 4 yards	5 to 10 yards	11+ yards	Negative air yards	1 to 10 air yards	11 to 20 air yards	21+ air yards

Play Bucket	**Adjusted EPA**
Strip sack	-3.71
Sack	-1.42
Rush big loss	-1.17
Rush small loss	-0.86
Rush no gain	-0.57
Screen under pressure	-0.44
Rush small gain	-0.11
Short pass under pressure	-0.07
Screen	-0.03
Deep pass under pressure	-0.02
Medium pass under pressure	0.14
Short pass	0.18
Deep pass	0.31
Scramble	0.36
Rush decent gain	0.38
Medium pass	0.44
Rush big gain	1.21

Relating Variation Within Positions to Outcomes

$$\text{Adj EPA} = \beta_0 + \beta_1(\text{Avg Normalized EDL Rtg}) + \beta_2(\text{Avg Normalized IDL Rtg})$$

Running the above linear regression, we determine the relationship between positional talent and outcomes.

- Adj EPA: Context-neutral, defensive-line-adjusted EPA for each play
- Avg Normalized EDL Rtg: the mean rating of EDLs for each play, normalized
- Avg Normalized IDL Rtg: the mean rating of IDLs for each play, normalized

β_1 represents the change in adjusted EPA associated with a one-standard-deviation increase in the average Madden rating of EDLs, holding IDLs constant
β_2 represents the change in adjusted EPA associated with a one-standard-deviation increase in the average Madden rating of IDLs, holding EDLs constant

Note: EPA is quoted for offense, so lower EPA means better defense

Relating Variation Within Positions to Outcomes

ALL FORMATIONS	β_1	β_2	$\beta_1 - \beta_2$
Against All Plays	-0.0159 *	-0.00970	0.0062
Against the Run	0.0000127	0.00771	0.00770
Against the Pass	-0.0278 **	-0.0205 **	0.0073

Significance Level Labels: 0.05 = * , 0.01 = **

- The regression coefficient in all plays for EDLs (β_1) is more negative than that for IDLs (β_2), indicating that variations in Exterior Defensive Linemen talent have a stronger impact on outcomes
  - Additionally, β_1 is significantly different from 0, while β_2 is not
- While increasing the skill level of both IDLs and EDLs has essentially no effect on run plays (as indicated by the coefficients that are approximately 0), both IDLs and EDLs have a significant impact on pass plays, with EDLs exhibiting a larger expected impact when their skill level is improved

Relating Variation Within Positions to Outcomes

- As previously seen, when considering all formations and personnel groupings, Exterior Defensive Linemen are more important
- However, the 2-2 formation (two EDLs, two IDLs) is widely used (42.1% of the plays in the dataset), and produces a different result
    - This regression indicates that Interior Defensive Linemen are more important, largely due to their significant impact against the run relative to Exterior Defensive Linemen

$$\text{Adj EPA} = \beta_0 + \beta_1(\text{Avg Normalized EDL Rtg}) + \beta_2(\text{Avg Normalized IDL Rtg})$$

2-2 FORMATION	β_1	β_2	$\beta_1 - \beta_2$
Against All Plays	-0.00593	-0.0134	0.00747
Against the Run	0.0149	-0.00828	0.0232
Against the Pass	-0.022 *	-0.0155	0.0065

Relating Variation Within Positions to Outcomes: *Conclusions*

By regressing outcomes on normalized positional ratings across the entire data set, we find that EDLs are slightly more important than IDLs.

When comparing run plays and pass plays, we found that neither the variation of talent in EDLs nor IDLs has a significant impact on run play outcomes. However, increased talent of both EDLs and IDLs does have an impact on pass play outcomes, with the impact of EDLs being more significant.

Contrary to the analysis that includes all formations, we found that when focusing solely on the 2-2 formation, IDLs are more important by a slight margin due to their variation having a greater influence on rushing outcomes.

Context-Specific Value

Combining these results with the probability of passing or rushing at a given game situation, we can model the value of increased talent at a defensive line position for each game situation.

- We model pass and run probabilities with a logistic regression, based on down, distance, score differential, and time remaining.

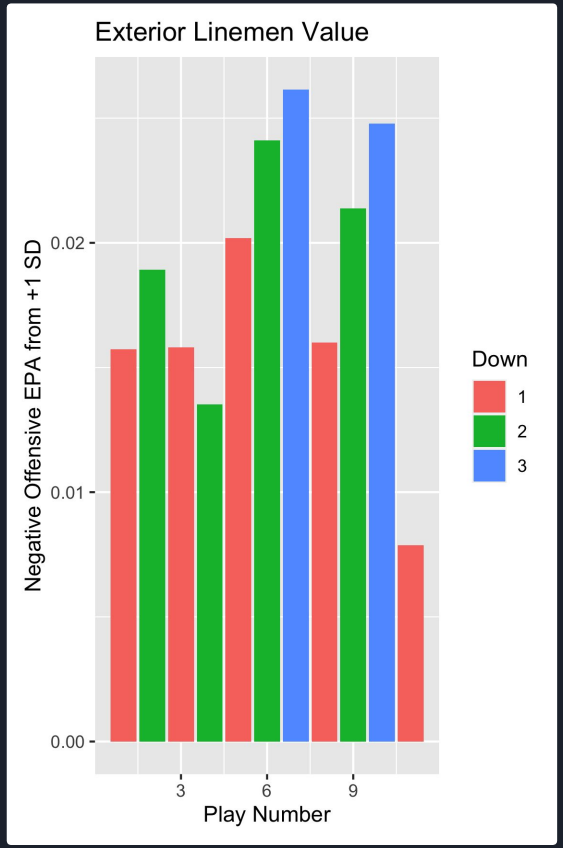
- Example: Kansas City Chiefs touchdown drive in 4th quarter of Super Bowl LIV
    - Kansas City was down 20-10 at the beginning of the drive with 8:53 remaining

Context-Specific Value Example

The barchart on the right displays the value of a one-standard-deviation increase in exterior defensive linemen talent in various situations along the Chiefs' drive.

- Y-axis: Negative of change in Adjusted EPA; larger values indicate better defensive outcomes

Play #	Down	To Go	Time
1	1	10	8:53
2	2	7	8:33
3	1	10	8:12
4	2	1	7:50
5	1	15	7:23
6	2	15	7:17
7	3	15	7:13
8	1	10	6:35
9	2	10	6:30
10	3	10	6:23
11	1	1	6:13



Limitations and Future Analysis

Limitations and Future Analysis

Analysis 1: Using a broader dataset to regress outcomes on specific positions

- We used positional groups (IDLs vs EDLs) because of the lack of stability in regression coefficients when regressing outcomes on specific positions (e.g. left EDL). More data would improve the stability of those coefficients, allowing us to extend this methodology to look granularly at specific positions.

Analysis 2: Modeling non-linear relationships between skill and outcomes

- The true relationship between skill and outcomes is likely non-linear. The linear model allows for generalizable conclusions, but in the context of one team's decisions, a non-linear model relating their specific players' skills to outcomes would be more valuable. This approach would also require a larger dataset.

Limitations and Future Analysis

Analysis 3: Accounting for confounding variables, such as quality of other players

- Our analysis suffers from potential confounding: highly rated defensive linemen may be correlated with high expenditure on the defensive line, leading to less expenditure (and thus lower quality) at other parts of the defense. With data on the other defensive players, as well as the opposing offense, we could account for this confounding effect.

Analysis 4: Performing a cost-benefit analysis of increasing player skill

- Our analysis helps quantify the benefit to a team of increasing skill by one standard deviation at a defensive line position. Combined with free agent market data, we could analyze the trade-off required to increase skill at a position, identifying the most efficient uses of team resources.

Citations

Competition Dataset:
(https://github.com/SportsInfoSolutions/AnalyticsChallenge2020)

Madden 20 Ratings:
(https://www.easports.com/madden-nfl/player-ratings)

Analysis and graphics using R + Photoshop